Journal of Multivariate Analysis 105 (2012) 368–379

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Self-consistent estimation of censored quantile regression

Limin Peng

Department of Biostatistics and Bioinformatics, Emory University, 1518 Clifton Rd. NE, Atlanta, 30322, USA

A R T I C L E   I N F O

Article history:
Received 11 June 2011
Available online 21 October 2011

AMS subject classifications:
62N01
62H99

Keywords:
Censoring
Martingale
Regression quantile
Self-consistency
Stochastic integral equation

A B S T R A C T

The principle of self-consistency has been employed to estimate regression quantile with randomly censored response. The asymptotic studies for this type of approach was established only recently, partly due to the complex forms of the current self-consistent estimators of censored regression quantiles. Of interest, how the self-consistent estimation of censored regression quantiles is connected to the alternative martingale-based approach still remains uncovered. In this paper, we propose a new formulation of self-consistent censored regression quantiles based on stochastic integral equations. The proposed representation of censored regression quantiles entails a clearly defined estimation procedure. More importantly, it greatly simplifies the theoretical investigations. We establish the large sample equivalence between the proposed self-consistent estimators and the existing estimator derived from martingale-based estimating equations. The connection between the new self-consistent estimation approach and the available self-consistent algorithms is also elaborated.

1. Introduction

Quantile regression [12] has arisen into a useful regression technique for survival data (i.e. time-to-event data). Compared to traditional survival regression methods, including the Cox proportional hazards model and the accelerated failure time (AFT) model, quantile regression can accommodate a more general relationship between an event time of interest and covariates while providing straightforward interpretations.

Let T and C denote time to event and time to censoring, and let $\boldsymbol{Z}$ denote a $p \times 1$ covariate vector with the first component set as 1. Without loss of generality, the censored regression quantile model investigated in this paper takes the form

$$Q_T(\tau|\boldsymbol{Z}) = \exp\{\boldsymbol{Z}^T \boldsymbol{\beta}_0(\tau)\}, \quad \tau \in (0, 1), \tag{1}$$

where $Q_T(\tau|\boldsymbol{Z}) \equiv \inf\{t : \Pr(T \le t|\boldsymbol{Z}) \ge \tau\}$ denotes the conditional quantile function of T given $\boldsymbol{Z}$ (with the same definition applied to any other random variable), and $\boldsymbol{\beta}_0(\tau)$ is a $p \times 1$ vector of unknown regression coefficients representing covariate effects on the τ-th quantile of $\log T$. Model (1) adopts the standard random censoring mechanism. That is, T and C are assumed to be independent conditional on $\boldsymbol{Z}$. Under this independent censoring assumption, the distribution of C is allowed to depend on $\boldsymbol{Z}$.

It is worth noting that much previous work on quantile regression with censored data cannot address the regression quantile problem defined above. For example, early efforts by Powell [20,21] require that all censoring times be known or fixed and so do the subsequent work by Fitzenberger [5], Buchinsky and Hahn [3], among others. Other methods, such as those by Ying et al. [25] and Honore et al. [7], demand unconditional independence between T and C, which is a stronger assumption than the standard random censorship, and thus cannot be applied here.

Portnoy [18] made the first attempt to tackle the censored regression quantile problem (1) by novelly employing the principle of self-consistency [4]. The principle of self-consistency here, in short, refers to an estimation scenario from

E-mail address: lpeng@sph.emory.edu.

0047-259X/$ – see front matter © 2012 Published by Elsevier Inc.
doi:10.1016/j.jmva.2011.10.005

equaling the estimator to an expression that contains the estimator itself. The estimator proposed in [18] reduces to the Kaplan–Meier estimator [9] in the one-sample case. The initial iterative algorithm was polished by Neocleous et al. [16] to a grid-based estimation procedure. We hereafter refer the estimator defined in [16] to as Portnoy's self-consistent estimator, denoted by $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$. Asymptotic studies on the self-consistent estimation of regression quantile were established only until recently by Portnoy and Lin [19], partly due to the complex representation of the existing estimators. Other subsequent work of [18] includes (but is not limited to) [15], which studied the monotonicity property associated with $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$, and [24], which relaxed the global linear assumption in model (1).

Peng and Huang [17] proposed an alternative approach to estimating $\boldsymbol{\beta}_0(\tau)$ in model (1) by utilizing the martingale structure of randomly censored data. A clearly defined grid-based algorithm was developed based on a set of monotone estimating equations. These estimating equations have appealing stochastic integral representations which greatly facilitate large-sample studies. Peng and Huang [17] derived the closed form of the limit process of their estimator, denoted by $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$. It was also shown that $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$ becomes a Nelson–Aalen type estimator [14,1] when there is no covariate. More recently, Huang [8] derived a grid-free estimation procedure based on a new concept of quantile calculus. The resulting estimator has the same distribution as Peng and Huang [17]'s estimator.

A brief introduction of the algorithms presented in [16,17] is provided in Section 2. Both of these approaches have been implemented for *R* in the contributed package *quantreg* [11]. Comprehensive numerical studies conducted by Koenker [10] demonstrated very similar empirical performance between $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$ and $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$. Though the asymptotic equivalence between these two estimators can be established in the one-sample case given the fact that the Kaplan–Meier estimator and the Nelson–Aalen estimator are equivalent in the large sample sense, their connection remains unknown in general regression settings.

In this paper, we develop a new framework to study the self-consistent estimation of model (1) with T subject to standard random censorship. With the principle of self-consistency generally perceived as an estimation strategy that defines an estimator as some function of the observed data and the estimator itself, and then utilizes such a relationship to obtain a consistent well-defined estimator, throughout this paper, the self-consistent estimation of model (1) refers to as the estimation of $\boldsymbol{\beta}_0(\cdot)$ which involves a use of self-consistency principle. Compared to previous work, the proposed method preserves the computational-ease feature, while providing a more direct approach to the asymptotic theory. Of note, the current self-consistent estimators are defined under the assumption that no censoring would occur below $\boldsymbol{Z}^T\boldsymbol{\beta}_0(\epsilon_l)$ for all $\boldsymbol{Z}$, where ϵ_l is a prespecified constant in $(0, 1)$. This assumption is not likely to incur serious concerns in real data analysis with ϵ_l selected small enough. Nevertheless, it is practically desirable to eliminate this restriction. In Section 3, we formulate censored regression quantiles based on stochastic integral equations derived from adopting self-consistency principle. The new representation of self-consistent censored regression quantiles entails clearly defined estimation and inference procedures, which avoid the artificial data constraint stated above. In Section 4, we show the asymptotic equivalence of the proposed self-consistent estimators to $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$. Therefore, we establish the uniform consistency and weak convergence of the new estimators with the closed-forms for the limit distributions derived. Furthermore, we elaborate the connection between the proposed estimators and Portnoy's self-consistent estimator. The results aid in understanding the close proximity between $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$ and $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$ observed in previous empirical studies. Monte-Carlo simulations reported in Section 5 confirm our theoretical findings.

It is important to note that the proposed framework for self-consistent estimation of censored quantile regression can be extended to survival settings with more complex censoring mechanisms. A few remarks are supplied in Section 6.

2. Two existing approaches for censored quantile regression

Define $\tilde{X} = T \wedge C$ and $\delta = I(T \le C)$, where $\wedge$ is the minimum operator and $I(\cdot)$ is the indicator function. The observed randomly censored data consist of n iid replicates of $(\tilde{X}, \delta, \boldsymbol{Z})$, denoted by $\{(\tilde{X}_i, \delta_i, \boldsymbol{Z}_i),\ i = 1, \ldots, n\}$. We define $X = \log \tilde{X}$ and accordingly $X_i = \log \tilde{X}_i$.

2.1. Portnoy's self-consistent approach

We here outline the grid-based algorithm presented in [16]. A grid of τ-values is defined as $0 < \tau_1 < \tau_2 < \cdots < \tau_M < 1$. Define $m(\boldsymbol{\beta}, i, k) = \max\{l : 1 \le l \le k-1, \boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau_l) < X_i \le \boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau_{l+1})\}$ if the set $\{l : 1 \le l \le k-1,\ \boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau_l) < X_i \le \boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau_{l+1})\}$ is not empty, and $m(\boldsymbol{\beta}, i, k) = k+1$ otherwise. By this definition, $m(\boldsymbol{\beta}, i, 0) = 1$.

Step 1. Compute $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$ by fitting the uncensored quantile regression with data $\{X_i, \delta_i, \boldsymbol{Z}_i\}_{i=1}^n$. It is assumed that all censored X_i's are above the hyperplane determined by $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$. Set $k = 1$.

Step 2. Given $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_l)$ $(l \le k)$, obtain $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_{k+1})$ by minimizing the following weighted check function:

$$\sum_{\delta_i=1} \rho_\tau(X_i - \boldsymbol{Z}_i^T\boldsymbol{b}) + \sum_{\delta_i=0}\Big\{\hat{w}_{k+1,\,i}\rho_\tau(X_i - \boldsymbol{Z}_i^T\boldsymbol{b}) + (1-\hat{w}_{k+1,\,i})\rho_\tau(X^* - \boldsymbol{Z}_i^T\boldsymbol{b})\Big\}, \tag{2}$$

where $\hat{w}_{k+1,i} = (\tau_{k+1} - \tau_{m(\widehat{\boldsymbol{\beta}}_{PSC},i,k)})/(1-\tau_{m(\widehat{\boldsymbol{\beta}}_{PSC},i,k)})$ if $m(\widehat{\boldsymbol{\beta}}_{PSC}, i, k) < k+1$ and 1 otherwise, $\rho_\tau(u) = u\{\tau - I(u < 0)\}$, and X^* is an extremely large value.

Step 3. Replace k by $k+1$.
Step 4. Repeat steps 2–3 until $k > M$ or only censored observations remain above $\boldsymbol{Z}_i^T \widehat{\boldsymbol{\beta}}_{PSC}(\tau_{k-1})$.

Remark 1. The censored regression quantiles studied in [19] slightly variate from those defined in [16]. First, Portnoy and Lin [19] defined both $\widehat{\boldsymbol{\beta}}(\tau_1)$ and $\widehat{\boldsymbol{\beta}}(\tau_2)$ by ordinary uncensored regression quantile methods. Second, Portnoy and Lin [19] used a linear interpolation to define the quantile crossing of a censored observation, while [16] essentially adopted the cadlag version of regression quantiles. Third, Portnoy [19] set $\hat{w}_{k+1,i}$ as 0 instead of 1 when $m(\widehat{\boldsymbol{\beta}}_{PSC}, i, k) = k+1$. Nevertheless, as pointed out by Portnoy and Lin [19], "either definition leads to the same asymptotic distribution".

Note that there may involve some ambiguity in determining $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$ when there is a censored X_i lying below the hyperplane $\boldsymbol{Z}_i^T \widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$. This complication tends to be more troublesome in bootstrap-based inferences, as the above phenomenon can occur more often in re-sampled datasets. A detailed discussion on this issue can be found in the Appendix of [18].

2.2. *Peng and Huang (2008) [17]'s approach*

The estimator proposed by Peng and Huang [17], $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$, is defined as a cadlag approximation to the solution of the estimating equation,

$$n^{1/2}\boldsymbol{S}_n^{(PH)}(\boldsymbol{\beta}, \tau) = 0,$$

where $\boldsymbol{S}_n^{(PH)}(\boldsymbol{\beta}, \tau) = n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i \left[N_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\} - \int_0^\tau Y_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}dH(u)\right]$ with $N_i(x) = I(X_i \le x, \delta_i = 1)$, $Y_i(x) = I(X_i \ge x)$ and $H(x) = -\log(1-x)$. A grid of τ-values, $\mathcal{G}$, is defined as $0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_M = \tau_U$, where τ_U is a deterministic constant subject to some identifiability constraint. Let $\|\mathcal{G}\|$ denote the size of $\mathcal{G}$, $\max_{k=0,\ldots,M}(\tau_{k+1} - \tau_k)$. Without further mentioning, $\mathcal{G}$ will be adopted throughout the rest of the paper, even for $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$. The algorithm for obtaining $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$ is as follows.

Step 1. Set $\exp\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(\tau_0)\} = 0$ for all i by the definition of $Q_T(\tau|\boldsymbol{Z})$. Set $k = 0$.
Step 2. Given $\exp\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(\tau_l)\} = 0$ for $l \le k$, obtain $\widehat{\boldsymbol{\beta}}_{PH}(\tau_{k+1})$ as the minimizer of the following L_1-type convex objective function:

$$l_{k+1}(\boldsymbol{h}) = \sum_{i=1}^n \left|\delta_i X_i - \delta_i \boldsymbol{h}^T\boldsymbol{Z}_i\right| + \left|X^* - \boldsymbol{h}^T\sum_{l=1}^n(-\delta_l\boldsymbol{Z}_l)\right| + \left|X^* - \boldsymbol{h}^T\sum_{r=1}^n\left(2\boldsymbol{Z}_r\sum_{l=0}^k I[X_r \ge \boldsymbol{Z}_r^T\widehat{\boldsymbol{\beta}}_{PH}(\tau_l)]\{H(\tau_{l+1}) - H(\tau_l)\}\right)\right|,$$

where X^* is an extremely large value.
Step 3. Replace k by $k+1$ and repeat step 2 until $k = M$ or no feasible solution can be found for minimizing $l_k(\boldsymbol{h})$.

3. The proposed formulation of self-consistent censored regression quantiles

Define $R_i(x) = I(X_i \le x, \ \delta_i = 0)$. Without considering covariates, Efron [4] suggested a self-consistent estimating equation for $F_{\log T}(t) \equiv \Pr(\log T \le t)$ given by

$$F_{\log T}(t) = n^{-1}\sum_{i=1}^n\left\{N_i(t) + R_i(t)\frac{F_{\log T}(t) - F_{\log T}(X_i)}{1 - F_{\log T}(X_i)}\right\}. \tag{3}$$

The right hand side (RHS) of Eq. (3) can be expressed as a stochastic integral. Specifically,

$$\begin{aligned}\text{RHS of (3)} &= n^{-1}\sum_{i=1}^n\left\{N_i(t) + R_i(t)\int_0^t \frac{F_{\log T}(t) - F_{\log T}(u)}{1 - F_{\log T}(u)}dR_i(u)\right\}\\ &= n^{-1}\sum_{i=1}^n\left[N_i(t) + R_i(t)\{1 - F_{\log T}(t)\}\int_0^t \frac{R_i(u)}{\{1 - F_{\log T}(u)\}^2}dF_{\log T}(u)\right],\end{aligned}$$

where the second equality follows by applying stochastic integral by parts assuming the continuity of $F_{\log T}(u)$. This renders an alternative form of the self-consistent estimating equation for $F_{\log T}(t)$,

$$F_{\log T}(t) = n^{-1}\sum_{i=1}^n\left[N_i(t) + R_i(t)\{1 - F_{\log T}(t)\}\int_0^t \frac{R_i(u)}{\{1 - F_{\log T}(u)\}^2}dF_{\log T}(u)\right]. \tag{4}$$

Interestingly, the self-consistent estimating Eq. (4), unlike Eq. (3), offers a natural extension to the general regression setting under model (1). Note that model (1) is equivalent to

$$Q_{\log T}(\tau|\boldsymbol{Z}) = \boldsymbol{Z}^T\boldsymbol{\beta}_0(\tau), \quad \tau \in (0, 1).$$

With t replaced by $\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)$, Eq. (4) evolves into an estimating equation for $\boldsymbol{\beta}_0(\tau)$:

$$n^{1/2}\boldsymbol{S}_n^{(SC)}(\boldsymbol{\beta},\tau)=0, \tag{5}$$

where $\boldsymbol{S}_n^{(SC)}(\boldsymbol{\beta},\tau)$ equals

$$n^{-1}\sum_{i=1}^{n}\boldsymbol{Z}_i\left[N_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}+R_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}(1-\tau)\int_0^{\tau}\frac{R_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}}{(1-u)^2}du-\tau\right].$$

The stochastic integral equation (5) provides a new formulation of censored regression quantiles driven by the use of self-consistency principle. It entails a simple and clearly-defined estimation procedure for $\boldsymbol{\beta}_0(\tau)$. The key idea is to approximate the solution to (5) by mimicking Euler's method for first-order differential equation. Specifically, let $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ denote a cadlag function, which only jumps at the τ-grid $\mathcal{G}$. Setting $\exp\{\boldsymbol{Z}_i\widehat{\boldsymbol{\beta}}_{SC}(0)\}=0$ for all i, one can obtain $\widehat{\boldsymbol{\beta}}_{SC}(\tau_{k+1})$ by sequentially solving the following equation for $\boldsymbol{b}$ for $k=0,\ldots,M-1$:

$$n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i\left(N_i(\boldsymbol{Z}_i^T\boldsymbol{b})+R_i(\boldsymbol{Z}_i^T\boldsymbol{b})\left[\sum_{l=0}^{k}R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau_l)\}\cdot\left(\frac{1-\tau_{k+1}}{1-\tau_{l+1}}-\frac{1-\tau_{k+1}}{1-\tau_l}\right)\right]-\tau_{k+1}\right)=0. \tag{6}$$

It is easy to see that Eq. (6) is a monotone estimating equation [6]. Simple algebraic manipulationsshow that the estimating function in (6) is the minus subgradient of the following weighted check function:

$$n^{-1/2}\left[\sum_{\delta_i=1}\rho_\tau(X_i-\boldsymbol{Z}_i^T\boldsymbol{b})+\sum_{\delta_i=0}\left\{\tilde{w}_{k+1,i}\rho_\tau(X_i-\boldsymbol{Z}_i^T\boldsymbol{b})+(1-\tilde{w}_{k+1,i})\rho_\tau(X^*-\boldsymbol{Z}_i^T\boldsymbol{b})\right\}\right], \tag{7}$$

where $\tilde{w}_{k+1,i}=\sum_{l=0}^{k}R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau_l)\}\left(\frac{1-\tau_{k+1}}{1-\tau_{l+1}}-\frac{1-\tau_{k+1}}{1-\tau_l}\right)$, $k=0,\ldots,M-1$.

Remark 2. As in [17], we set $\exp\{\boldsymbol{Z}_i\widehat{\boldsymbol{\beta}}_{SC}(0)\}=0$ for all i because $Q_T(0|\boldsymbol{Z}_i)=0$ according to the definition of conditional quantile (i.e. $Q_T(\tau|\boldsymbol{Z})\equiv\inf\{t:\ \Pr(T\le t|\boldsymbol{Z}\ge\tau)\}$). Similarly, the justification of this boundary constraint requires the smoothness of $\boldsymbol{\beta}_0(\tau)$ and positive density of T everywhere between 0 and the τ_U-th quantile, implied by the regularity conditions C2(a) and C3(a) in [17] respectively. Note that these assumptions would be violated when the support of T is bounded away from 0, i.e. $[c,+\infty)$ with $c>0$. In such a case, an easy adaptation is to transform T by $T-c$ if c is known. When c is unknown, a practical remedy may be to fit model (1) to $\max(T-c',0)$, where c' may be chosen as the observed lower bound of T's support. This is expected to result in only minimal deviations from the original regression quantiles of interest when c' is close to c.

As shown above, the self-consistent estimator $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ is well defined without imposing additional constraints on the support of C. This estimator can be easily computed in a sequential manner like Portnoy's self-consistent estimator. Another appealing feature is that $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ is formulated based on a stochastic integral equation which provides an easy platform for conducting large-sample studies. This enables in-depth investigations on the connections between $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ and existing estimators, including $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$.

4. Asymptotic results

Large sample studies for $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ is greatly facilitated by the stochastic integral equation representation of (5). In Section 4.1, we first show the asymptotic equivalence between $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$. This result immediately implies the uniform consistency and weak convergence of the new self-consistent estimator $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, provided the asymptotic properties established for $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$ in [17]. Since the closed-form limit distribution of $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$ reduces to that of the Kaplan–Meier estimator in the one-sample case, the same property applies to $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$.

In Section 4.2, we further exploit the relationship between $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, which ultimately uncovers the underlying connection between $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$. Theoretical work conducted here also suggests several variants of $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$.

4.1. *Asymptotic equivalence between $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$*

The result is formally stated in the following theorem:

Theorem 1. *Suppose model (1) holds. Assuming conditions required by Theorem 2 of* [17] *and condition* (D1) *in Appendix are satisfied,*

$$\sup_{\tau\in[\nu,\tau_U]}\|n^{1/2}\{\widehat{\boldsymbol{\beta}}_{PH}(\tau)-\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}\|\to_p 0,$$

for any $\nu\in(0,\tau_U)$.

Our proof of Theorem 1 involves a variant of estimating Eq. (5) given by

$$n^{1/2}\boldsymbol{S}_n^{(OSC)}(\boldsymbol{\beta},\tau)=0,\quad \tau\in(0,\tau_U],\tag{8}$$

where

$$\boldsymbol{S}_n^{(OSC)}(\boldsymbol{\beta},\tau)=n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i\left[N_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}+(1-\tau)\int_0^\tau \frac{R_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(s)\}}{(1-s)^2}ds-\tau\right].$$

A self-consistent estimator of $\boldsymbol{\beta}_0(\tau)$, $\widehat{\boldsymbol{\beta}}_{OSC}(\tau)$, can be derived based on Eq. (8) in the same fashion that $\widehat{\boldsymbol{\beta}}^{(SC)}(\tau)$ is defined based on Eq. (5). Specifically, $\widehat{\boldsymbol{\beta}}_{OSC}(\tau)$ is a cadlag process obtained as follows: (i) set $\exp\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(\tau_0)\}=0$ for all i; (ii) for $k=0,\ldots,M-1$, sequentially compute $\widehat{\boldsymbol{\beta}}_{OSC}(\tau_{k+1})$ as the minimizer of the L_1-type convex objective function, $\ell_{k+1}(\boldsymbol{h})=$

$$\sum_{i=1}^n\left|\delta_iX_i-\delta_i\boldsymbol{h}^T\boldsymbol{Z}_i\right|+\left|X^*-\boldsymbol{h}^T\sum_{l=1}^n(-\delta_l\boldsymbol{Z}_l)\right|+\left|X^*-2Z_r\sum_{r+1}^n(\tau_{k+1}-\tilde{w}_{k+1,r})\right|.$$

In our arguments, $\widehat{\boldsymbol{\beta}}_{OSC}(\tau)$ serves as a bridge to connect $\widehat{\boldsymbol{\beta}}_{SC}(\tau)$ and $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$. By the Lemmas 2 and 3 stated below, we can respectively show the asymptotic equivalence between $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$ and $\widehat{\boldsymbol{\beta}}_{OSC}(\tau)$ and that between $\widehat{\boldsymbol{\beta}}_{SC}(\tau)$ and $\widehat{\boldsymbol{\beta}}_{OSC}(\tau)$, thereby obtaining the result in Theorem 1.

Lemma 2. *Suppose that model (1) holds. Let $\tilde{\boldsymbol{\beta}}(\cdot)$ and $\check{\boldsymbol{\beta}}(\cdot)$ be cadlag processes that only jump at $0<\tau_1<\tau_2<\cdots<\tau_M<\tau_U$. Assuming that conditions required by Theorem 2 of [17] and condition (D1) are satisfied, if $n^{1/2}\boldsymbol{S}_n^{(u)}(\tilde{\boldsymbol{\beta}},\tau)\stackrel{a}{=}0$ and $n^{1/2}\boldsymbol{S}_n^{(u)}(\check{\boldsymbol{\beta}},\tau)\stackrel{a}{=}0$, then*

$$\sup_{\tau\in[\nu,\tau_U]}\|\tilde{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}_0(\tau)\|\to_p 0,\qquad \sup_{\tau\in[\nu,\tau_U]}\|\check{\boldsymbol{\beta}}(\tau)-\boldsymbol{\beta}_0(\tau)\|\to_p 0,$$

and

$$\sup_{\tau\in[\nu,\tau_U]}\|n^{1/2}\{\tilde{\boldsymbol{\beta}}(\tau)-\check{\boldsymbol{\beta}}(\tau)\}\|\to_p 0$$

for any $\nu\in(0,\tau_U)$, where the superscript u can be "PH", "SC", and "OSC". Here and in the sequel, $\stackrel{a}{=}$ means asymptotic equivalence uniformly in $\tau\in(0,\tau_U]$.

Remark 3. Lemma 2 states that a deviation of $o(1)$ (uniformly in τ) in the estimating function $n^{1/2}\boldsymbol{S}_n^{(PH)}(\cdot)$, $n^{1/2}\boldsymbol{S}_n^{(SC)}(\cdot)$, or $n^{1/2}\boldsymbol{S}_n^{(OSC)}(\cdot)$ does not affect the uniform consistency and only change the limit process of the resulting estimator by $o(n^{-1/2})$.

Lemma 3. *Suppose that model (1) holds. Assuming that conditions required by Theorem 2 of [17] and condition (D1) are satisfied,*

$$\sup_{\tau\in(0,\tau_U]}\left\|n^{-1/2}\sum_{i=1}^n\boldsymbol{Z}_iI\{X_i>\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}\int_0^\tau\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(s)\}}{(1-s)^2}ds\right\|\to_p 0.\tag{9}$$

The technical proofs of Lemmas 2 and 3, and Theorem 1 are relegated to Appendix.

4.2. Connection between $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$

We start with an introduction of some new notation. Define $A_i(\boldsymbol{\beta},\tau)=\{u:0\le u<\tau,\boldsymbol{Z}_i^T\boldsymbol{\beta}(u-)\le X_i\le\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}$ and $\psi_i(\boldsymbol{\beta},\tau)=\sup\{A_i(\boldsymbol{\beta},\tau)\}\cdot I(A_i(\boldsymbol{\beta},\tau)\text{ is not empty})+\tau\cdot I(A_i(\boldsymbol{\beta},\tau)\text{is empty})$. It is easy to see that $\psi_i(\widehat{\boldsymbol{\beta}}_s,\tau_{k+1})=\tau_{m(\widehat{\boldsymbol{\beta}}_s,i,k)}$, where the subscript s can be "PH", "SC", "OSC", or "PSC". It is easy to see that $\psi_i(\boldsymbol{\beta},\tau)$ stands for the largest τ at which $\boldsymbol{Z}_i^T\boldsymbol{\beta}(\cdot)$ crosses X_i.

To exploit the connection between $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$, it is worthwhile to compare the objective functions adopted in the corresponding algorithms, given by (2) and (7) respectively. It is observed that both objective functions take a form of weighted sum of check functions, while the major distinction lies with the weight definitions. It is interesting to note that the weight distinction may vanish if $\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ are nondecreasing for all i. This is because with nondecreasing $\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$,

$$n^{1/2}\boldsymbol{S}_n^{(SC)}(\widehat{\boldsymbol{\beta}}_{SC},\tau)=n^{-1/2}\sum_{i=1}^n\boldsymbol{Z}_i\left[N_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}+R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)\int_{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}^\tau\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(u)\}}{(1-u)^2}du-\tau\right]\tag{10}$$

$$=n^{-1/2}\sum_{i=1}^n\boldsymbol{Z}_i\left[N_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}+R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}\frac{\tau-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}{1-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}-\tau\right],\tag{11}$$

and the estimating function in (11) renders weights, $\frac{\tau-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}{1-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}$, that take the same form as those for $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$.

However, the procedure to obtain $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ does not ensure the increasingness of $\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ but its uniform convergence to $\boldsymbol{\beta}_0(\cdot)$ which satisfies the monotonicity requirement on $\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\cdot)$ for all i. In Lemma 4, we claim that the weight contributions from τ less than $\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)$ in (7) are negligible even when $\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ lacks the monotonicity. Lemma 4 is stated below with proof given in Appendix.

Lemma 4. *Suppose model* (1) *holds. Assuming that conditions required by Theorem 2 of* [17] *and condition* (D1) *are satisfied,*

$$\sup_{\tau\in(0,\tau_U]}\left\|n^{-1/2}\sum_{i=1}^n \boldsymbol{Z}_i R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)\int_0^{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\right\|\to_p 0.$$

It immediately follows from Lemma 4, the definition of $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, and (11) that

$$n^{1/2}\boldsymbol{S}_n^{(MSC)}(\widehat{\boldsymbol{\beta}}_{SC},\tau)\overset{a}{=}0, \tag{12}$$

where

$$\boldsymbol{S}_n^{(MSC)}(\boldsymbol{\beta},\tau)=n^{-1}\sum_{i=1}^n\boldsymbol{Z}_i\left[N_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}+R_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}\frac{\tau-\psi_i(\boldsymbol{\beta},\tau)}{1-\psi_i(\boldsymbol{\beta},\tau)}-\tau\right].$$

Remark 4. The form of $\boldsymbol{S}_n^{(MSC)}(\boldsymbol{\beta},\tau)$ bears some similarity with that of McKeague et al. [13]'s estimating function for median regression. Unlike in [13], we estimate $\{\tau-\psi_i(\boldsymbol{\beta},\tau)\}/\{1-\psi_i(\boldsymbol{\beta},\tau)\}$ self-consistently based on the assumed quantile regression model (1) rather than imposing a separate model of T given $\boldsymbol{Z}$.

Eq. (12) induces another variant of $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$, defined as a cadlag solution to

$$n^{1/2}\boldsymbol{S}_n^{(MSC)}(\boldsymbol{\beta},\tau)=0. \tag{13}$$

More specifically, $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$ jumps only on the grid $\mathcal{G}$ with $\exp\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{MSC}(0)\}=0$ for all i and $\widehat{\boldsymbol{\beta}}_{MSC}(\tau_{k+1})$, given $\{\widehat{\boldsymbol{\beta}}_{MSC}(\tau_l)\}_{l=0}^k$, is obtained as the solution to

$$n^{-1/2}\sum_{i=1}^n\boldsymbol{Z}_i\left[N_i(\boldsymbol{Z}_i^T\boldsymbol{b})+R_i(\boldsymbol{Z}_i^T\boldsymbol{b})\left(\frac{\tau_{k+1}-\tau_{m(\widehat{\boldsymbol{\beta}}_{MSC},i,k)}}{1-\tau_{m(\widehat{\boldsymbol{\beta}}_{MSC},i,k)}}\right)-\tau\right]=0. \tag{14}$$

Given $\|\mathcal{G}\|=o(n^{-1/2})$, (12) holds with $\widehat{\boldsymbol{\beta}}_{SC}$ replaced by $\widehat{\boldsymbol{\beta}}_{MSC}$.

Note that $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$ greatly resembles Portnoy's estimator $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$ by noting that the estimating function in (14) is the minus subgradient of the following weighted check function,

$$n^{-1/2}\left[\sum_{\delta_i=1}\rho_\tau(X_i-\boldsymbol{Z}_i^T\boldsymbol{b})+\sum_{\delta_i=0}\left\{w_{k+1,\,i}^*\rho_\tau(X_i-\boldsymbol{Z}_i^T\boldsymbol{b})+(1-w_{k+1,\,i}^*)\rho_\tau(X^*-\boldsymbol{Z}_i^T\boldsymbol{b})\right\}\right],$$

where $w_{k+1,\,i}^*=\{\tau_{k+1}-\tau_{m(\widehat{\boldsymbol{\beta}}_{MSC},i,k)}\}/\{1-\tau_{m(\widehat{\boldsymbol{\beta}}_{MSC},i,k)}\}$, $k=0,\ldots,M-1$. The weights for $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$, $\hat{w}_{k,i}$, and the weights for $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$, $w_{k,i}^*$, have nearly identical definitions except for the weights used for computing $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$ and $\widehat{\boldsymbol{\beta}}_{MSC}(\tau_1)$. The algorithm in [18] or [16] forces all $\hat{w}_{1,i}$'s equal to 1 when obtaining $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$ by uncensored regression quantiles. This may incur some complication if there are censored X_i's less than $\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$. In contrast, the weights for $\widehat{\boldsymbol{\beta}}_{MSC}(\tau_1)$, $w_{1,i}^*$'s, are set as 0 justified on the basis of estimating Eq. (14). By the similarity and the difference stated above, $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$ may be viewed as a modified version of $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$, which avoids the somewhat ad-hoc restriction of $C_i<\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau_1)$ as well as the associated algorithmic issues.

Asymptotic studies for $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$ are similar to those for $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$ presented in [19], involving the main difficulty in analyzing the term that involves $\{\tau-\psi_i(\boldsymbol{\beta},\tau)\}/\{1-\psi_i(\boldsymbol{\beta},\tau)\}$. Provided the uniform consistency of $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$, we can write the "difference in τ" term in $\boldsymbol{S}_n^{(MSC)}(\boldsymbol{\beta},\tau)$, namely

$$n^{-1}\sum_{i=1}^n\boldsymbol{Z}_iR_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}\frac{\tau-\psi_i(\boldsymbol{\beta},\tau)}{1-\psi_i(\boldsymbol{\beta},\tau)},$$

as a Riemann sum by adapting Steps 1–6 in the proof for Theorem 3.1 of [19]. Using similar arguments to their Step 7, we can show that a $o(1)$ (uniformly in τ) deviation in $n^{1/2}\boldsymbol{S}_n^{(MSC)}(\boldsymbol{\beta},\tau)$ has little impact on the limit process of the solution to the resulting equation. Note that Eq. (12) reveals that $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$ are both solutions to the equation $n^{1/2}\boldsymbol{S}_n^{(MSC)}(\boldsymbol{\beta},\tau)\overset{a}{=}0$. The asymptotic equivalence between the proposed self-consistent estimator, $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, and the modified Portnoy's estimator, $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$, then follows from this fact.

Table 1

Empirical biases $\times 10^3$ (Bias) and empirical variances $\times 10^3$ (Var) of $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$, $\widehat{\boldsymbol{\beta}}_{SC}(\tau)$, $\widehat{\boldsymbol{\beta}}_{OSC}(\tau)$, $\widehat{\boldsymbol{\beta}}_{MSC}(\tau)$, and $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$, with C1, C2, and C3 denoting the first, the second, and the third component of the corresponding estimator respectively.

τ		$\widehat{\boldsymbol{\beta}}_{PH}$		$\widehat{\boldsymbol{\beta}}_{OSC}$		$\widehat{\boldsymbol{\beta}}_{SC}$		$\widehat{\boldsymbol{\beta}}_{MSC}$		$\widehat{\boldsymbol{\beta}}_{PSC}$	
		Bias	Var	Bias	Var	Bias	Var	Bias	Var	Bias	Var
(I) Log-linear model with iid errors											
0.1	C1	3	250	12	251	12	251	12	252	71	272
	C2	0	588	6	594	6	594	5	597	2	635
	C3	30	188	33	188	34	189	34	188	33	200
0.3	C1	2	87	15	88	15	88	15	88	39	89
	C2	24	204	27	206	27	206	27	206	28	206
	C3	1	70	1	70	1	70	1	70	4	72
0.5	C1	2	55	13	54	13	54	13	54	32	55
	C2	22	139	24	140	24	140	25	140	25	137
	C3	5	49	4	50	4	50	4	50	3	51
0.7	C1	23	54	2	53	1	53	1	53	21	53
	C2	12	131	14	130	14	130	15	131	8	130
	C3	10	46	9	45	9	45	9	45	3	43
(II) Log-linear model with heteroscedastic errors											
0.1	C1	13	88	9	88	9	88	9	88	26	92
	C2	1	211	3	212	3	212	3	212	4	223
	C3	1	74	0	74	0	74	0	74	2	75
0.3	C1	9	49	1	49	1	49	1	50	17	50
	C2	1	114	2	115	2	115	2	115	1	117
	C3	5	43	4	43	4	43	4	43	4	43
0.5	C1	29	50	12	49	12	49	12	49	5	49
	C2	18	117	17	116	17	117	17	117	19	114
	C3	8	40	8	40	8	40	8	40	5	40
0.7	C1	43	68	14	68	14	68	15	68	7	65
	C2	10	156	5	156	4	156	4	157	6	152
	C3	14	51	14	50	15	50	15	50	12	48

5. Simulations

Simulation studies have been conducted to compare Peng and Huang's estimator $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$ and the four types of self-consistent estimators, including $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, $\widehat{\boldsymbol{\beta}}_{OSC}(\cdot)$, $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$, and $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$. For brevity of presentation, we only report results from two scenarios: (I) log-linear model with iid errors; (ii) log-linear model with heteroscedastic errors. The configurations follow those used in [17] with 25% censoring. Specifically, for scenario (I), T follows the model $\log T = 0.5Z_1 - 0.5Z_2 + \epsilon$, where $\epsilon \sim$ extreme value distribution, $Z_1 \sim \mathit{Unif}(0, 1)$, and $Z_2 \sim \mathit{Bernoulli}(.5)$. The censoring distribution is $\mathit{Unif}(0.1I(Z_2 = 1), 3.8)$. For scenario (II), T is generated from the model $\log T = 0.5Z_1 - 0.5Z_2\xi + \epsilon$, where $\xi \sim$ exponential(1), $\epsilon \sim N(0, 1)$, $Z_1 \sim \mathit{Unif}(0, 1)$, and $Z_2 \sim \mathit{Bernoulli}(.5)$. The censoring time $C \sim \mathit{Unif}(0.3I(Z_2 = 1), 5.2)$. Under each configuration, we generate 1000 simulated datasets of sample size $n = 200$.

In Table 1, we present the empirical biases (Bias) and empirical variances (Var) of all five estimators under comparison. As expected, all estimators have small empirical biases, and their empirical variances are very similar, especially those of $\widehat{\boldsymbol{\beta}}_{OSC}(\tau)$, $\widehat{\boldsymbol{\beta}}_{SC}(\tau)$ and $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$.

We also examine the difference among these estimators based on each simulated dataset. Selecting $\widehat{\boldsymbol{\beta}}_{SC}(\tau)$ as the reference, we present in Table 2 the empirical 25th percentiles (Diff25) and empirical 75th percentile (Diff75) of $\widehat{\boldsymbol{\beta}}_{PH}(\tau) - \widehat{\boldsymbol{\beta}}_{SC}(\tau)$, $\widehat{\boldsymbol{\beta}}_{OSC}(\tau) - \widehat{\boldsymbol{\beta}}_{SC}(\tau)$, $\widehat{\boldsymbol{\beta}}_{MSC}(\tau) - \widehat{\boldsymbol{\beta}}_{SC}(\tau)$, and $\widehat{\boldsymbol{\beta}}_{PSC}(\tau) - \widehat{\boldsymbol{\beta}}_{SC}(\tau)$. Results in Table 2 confirm the observation from Table 1 that the three new self-consistent estimator, $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, $\widehat{\boldsymbol{\beta}}_{OSC}(\cdot)$, and $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$, are in close proximity; in over 50% of simulations they coincide with each other. It is interesting to note that Portnoy's estimator $\widehat{\boldsymbol{\beta}}_{PSC}(\tau)$ seems to have a relatively larger deviation from $\widehat{\boldsymbol{\beta}}_{SC}(\tau)$ than does $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$ at small τ's. This is likely due to its special treatment on $\widehat{\boldsymbol{\beta}}_{PSC}(\tau_1)$. The deviation of Peng and Huang's estimator $\widehat{\boldsymbol{\beta}}_{PH}(\tau)$ from the three new self-consistent estimators appears to rise as τ increases. This phenomenon is not surprising and shares the same spirit as the increasing cumulative error of Euler's solution to an ordinary differential equation. In summary, our simulation results provide empirical evidence for the asymptotic equivalence among $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$, $\widehat{\boldsymbol{\beta}}_{SC}(\cdot)$, $\widehat{\boldsymbol{\beta}}_{OSC}(\cdot)$, and $\widehat{\boldsymbol{\beta}}_{MSC}(\cdot)$, and their close connections with $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$, in addition to the theoretical arguments given in Section 3.

6. Remarks

The principle of self-consistency has been widely adopted in survival analysis as an intuitive way to handle missing information due to censoring and/or truncation. Examples include estimating survival function with randomly censored

Table 2

Empirical 25th percentiles $\times 10^3$ (DIFF25) and empirical 75th percentiles $\times 10^3$ (DIFF75) of $\widehat{\boldsymbol{\beta}}_{PH}-\widehat{\boldsymbol{\beta}}_{SC}$, $\widehat{\boldsymbol{\beta}}_{OSC}-\widehat{\boldsymbol{\beta}}_{SC}$, $\widehat{\boldsymbol{\beta}}_{MSC}-\widehat{\boldsymbol{\beta}}_{SC}$, and $\widehat{\boldsymbol{\beta}}_{PSC}-\widehat{\boldsymbol{\beta}}_{SC}$, with C1, C2, and C3 denoting the first, the second, and the third component of the corresponding quantity respectively.

τ		$\widehat{\boldsymbol{\beta}}_{PH}-\widehat{\boldsymbol{\beta}}_{SC}$		$\widehat{\boldsymbol{\beta}}_{OSC}-\widehat{\boldsymbol{\beta}}_{SC}$		$\widehat{\boldsymbol{\beta}}_{MSC}-\widehat{\boldsymbol{\beta}}_{SC}$		$\widehat{\boldsymbol{\beta}}_{PSC}-\widehat{\boldsymbol{\beta}}_{SC}$	
		Diff25	Diff75	Diff25	Diff75	Diff25	Diff75	Diff25	Diff75
(I) Log-linear model with iid errors									
0.1	C1	0	0	0	0	0	0	−63	0
	C2	0	0	0	0	0	0	0	0
	C3	0	0	0	0	0	0	−7	4
0.3	C1	0	5	0	0	0	0	−33	0
	C2	0	0	0	0	0	0	0	0
	C3	0	0	0	0	0	0	−1	9
0.5	C1	0	17	0	0	0	0	−24	0
	C2	0	0	0	0	0	0	0	0
	C3	−6	6	0	0	0	0	−3	6
0.7	C1	0	39	0	0	0	0	−23	0
	C2	−18	12	0	0	0	0	−1	0
	C3	−13	14	0	0	0	0	−2	10
(II) Log-linear model with heteroscedastic errors									
0.1	C1	0	0	0	0	0	0	−48	0
	C2	0	0	0	0	0	0	−11	4
	C3	0	0	0	0	0	0	−14	9
0.3	C1	0	2	0	0	0	0	−23	0
	C2	0	0	0	0	0	0	−3	4
	C3	0	0	0	0	0	0	−6	6
0.5	C1	0	23	0	0	0	0	−24	0
	C2	−4	0	0	0	0	0	0	2
	C3	−7	7	0	0	0	0	−4	10
0.7	C1	0	48	0	0	0	0	−26	0
	C2	−27	12	0	0	0	0	−4	0
	C3	−14	16	0	0	0	0	−6	7

data [4], doubly censored data [22], and interval-censored data [23]. In this paper, we present a new representation of self-consistent regression quantiles with randomly censored data based on stochastic integral equations. Such a formulation allows us to carry out self-consistent estimation in a sequential manner, while facilitating asymptotic studies. The new framework can also be extended to other more complex settings where the principle of self-consistency is applicable.

The proposed self-consistent estimator is shown to be asymptotically equivalent to Peng and Huang [17]'s estimator derived based on the martingale structure of the data. This result may be viewed as an extension of Kaplan–Meier and Nelson–Aalen equivalence in the one-sample case to the quantile regression setting. The techniques adopted in this work are conceptually simple and have good adaptability to other self-consistent estimation settings.

It is worth emphasizing that we adopt a grid size of $o(n^{-1/2})$ for all estimators considered in the paper, including Portnoy's estimator $\widehat{\boldsymbol{\beta}}_{PSC}(\cdot)$. It is unclear to us whether the theoretical results can hold with a less finer τ-grid. The work by Neocleous and Portnoy [15] suggests that this may be possible in the absence of censoring along with a benefit of guaranteed monotonicity of estimated quantile functions. Investigations along this direction for censored quantile regression may merit future research but are beyond the scope of this paper.

Acknowledgments

The author thanks Professor Stephen Portnoy for his helpful comments on this work. This research was supported by the US National Science Foundation grant DMS-1007660.

Appendix

Define $F_X(x|\boldsymbol{Z}) = \mathrm{pr}(X \le x|\boldsymbol{Z})$, $F_{X,1}(x|\boldsymbol{Z}) = \mathrm{pr}(X \le x, \delta = 1|\boldsymbol{Z})$, $F_{\log T}(x|\boldsymbol{Z}) = \mathrm{pr}(\log T \le x|\boldsymbol{Z})$, $f_X(x|\boldsymbol{Z}) = dF_X(x|\boldsymbol{Z})/dx$, $f_{X,1}(x|\boldsymbol{Z}) = dF_{X,1}(x|\boldsymbol{Z})/dx$, and $f_{\log T}(x|\boldsymbol{Z}) = dF_{\log T}(x|\boldsymbol{Z})/dx$. Let $\mathcal{Z}$ denote the domain of $\boldsymbol{Z}$.

Regularity conditions include conditions (C1)–(C6) in [17], and

(D1): (i) $\sup_{x,\boldsymbol{z}\in\mathcal{Z}} f_X(x|\boldsymbol{z})$ is bounded; (ii) $\sup_{x,\boldsymbol{z}\in\mathcal{Z}} f_{\log T}(x|\boldsymbol{z})$ is bounded; (iii) Each component of $E[\boldsymbol{Z}^{\otimes 2} f_X(\boldsymbol{Z}^T\boldsymbol{b}|\boldsymbol{Z})]$ $(E[\boldsymbol{Z}^{\otimes 2} f_{X,1}(\boldsymbol{Z}^T\boldsymbol{b}|\boldsymbol{Z})])^{-1}$ is uniformly bounded in $\boldsymbol{b} \in \mathcal{B}(d_0)$, where $\mathcal{B}(d_0)$ is the same as $\mathcal{B}(d_0)$ defined in [17].

Proof of Lemma 2. In case that u represents "PH", Lemma 1 follows immediately from the proofs of Theorems 1–2 in [17]. Very similar arguments can be applied to prove the case with u representing "SC", and "OSC" and hence are omitted here. □

Proof of Lemma 3. First, it is easy to note that the LHS of (9) is bounded above by $c_1 n^{-1/2}\sum_{i=1}^n I\{X_i > \boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}\int_0^\tau \frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(u)\}}{(1-u)^2}du$, and thus by

$$c_1\cdot\int_0^\tau\left[n^{-1/2}\sum_{i=1}^n\frac{I\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(\tau)<X_i\le \boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(u)\}}{(1-u)^2}\right]du,$$

where c_1 is the upper bound for $\|\boldsymbol{Z}\|$. Then (9) holds if

$$\sup_{0<u<\tau\le\tau_U} n^{-1/2}\sum_{i=1}^n\frac{I\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(\tau)<X_i\le \boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(u)\}}{(1-u)^2}=o_p(1). \tag{15}$$

Let $\varphi(\boldsymbol{\beta},u,\tau)=\mathrm{pr}\{\boldsymbol{Z}^T\boldsymbol{\beta}(\tau)<X\le\boldsymbol{Z}^T\boldsymbol{\beta}(u)\}$. To prove (15), it suffices to show that $\sup_{0<u<\tau\le\tau_U}\varphi(\widehat{\boldsymbol{\beta}}_{OSC},u,\tau)=o_p(1)$. Define $\boldsymbol{\mu}(\boldsymbol{b})=E\{\boldsymbol{Z}I(X\le\boldsymbol{Z}^T\boldsymbol{b})\delta\}$, $\boldsymbol{\phi}(\boldsymbol{b})=E\{\boldsymbol{Z}I(X\le\boldsymbol{Z}^T\boldsymbol{b})\}$, and $\phi_1(\boldsymbol{b})=E\{I(X\le\boldsymbol{Z}^T\boldsymbol{b})\}$. By the proof for Theorem 2 in [17] and the asymptotic equivalence between $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$ and $\widehat{\boldsymbol{\beta}}_{OSC}(\cdot)$, we have $\sup_{\tau\in[0,\tau_U]}\|\boldsymbol{\mu}\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}-\boldsymbol{\mu}\{\boldsymbol{\beta}_0(\tau)\}\|\to_p 0$. By condition D1 (iii), we then have $\sup_{\tau\in[0,\tau_U]}\|\boldsymbol{\phi}\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}-\boldsymbol{\phi}\{\boldsymbol{\beta}_0(\tau)\}\|\to_p 0$.

For any $\vartheta>0$, we can find some ν_ϑ such that $\sup_{\tau\in[0,\nu_\vartheta]}|\phi_1\{\boldsymbol{\beta}_0(\tau)\}|\le\vartheta/8$ because $\phi_1\{\boldsymbol{\beta}_0(0)\}=0$ and $\phi_1\{\boldsymbol{\beta}_0(\tau)\}$ is Lipschitz-continuous in τ. Given the uniform convergence of $\boldsymbol{\phi}\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}$ to $\boldsymbol{\phi}\{\boldsymbol{\beta}_0(\tau)\}$, there exists $N_{\vartheta,\xi,1}>0$ such that for $n\ge N_{\vartheta,\xi,1}$,

$$\mathrm{pr}\left(\sup_{\tau\in[0,\tau_U]}|\phi_1\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}-\phi_1\{\boldsymbol{\beta}_0(\tau)\}|>\vartheta/8\right)<\xi/3.$$

When $\sup_{\tau\in[0,\tau_U]}|\phi_1\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}-\phi_1\{\boldsymbol{\beta}_0(\tau)\}|\le\vartheta/8$, we have

$$\sup_{0<u<\tau\le\tau_U,u<\nu_\vartheta}|\varphi\{\widehat{\boldsymbol{\beta}}_{OSC},u,\tau\}|\le\sup_{\tau\in[0,\nu_\vartheta]}|\phi_1\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}|<\vartheta/4.$$

Since $\sup_{\tau\in[\nu,\tau_L]}\|\widehat{\boldsymbol{\beta}}_{OSC}(\tau)-\boldsymbol{\beta}_0(\tau)\|\to_p 0$ for any $\nu\in(0,\tau_U)$ by Theorem 2 of [17], there exists $N_{\vartheta,\xi,2}$ such that for $n\ge N_{\vartheta,\xi,2}$,

$$\mathrm{pr}\left(\sup_{\tau\in[\nu_\vartheta,\tau_U]}\|\widehat{\boldsymbol{\beta}}_{OSC}(\tau)-\boldsymbol{\beta}_0(\tau)\|>\frac{\vartheta}{2c_2}\right)<\xi/3,$$

where c_2 is a positive constant that bounds $\sup_{x,\boldsymbol{z}}f_X(x|\boldsymbol{z})$ from above and its existence is guaranteed by condition D1 (i). Note that $\sup_{\tau\in[\nu_\vartheta,\tau_U]}\|\widehat{\boldsymbol{\beta}}_{OSC}(\tau)-\boldsymbol{\beta}_0(\tau)\|\le\vartheta/(2c_2)$ implies that

$$\begin{aligned}\sup_{\nu_\vartheta\le u<\tau\le\tau_U}\varphi(\widehat{\boldsymbol{\beta}}_{OSC},u,\tau)&\le\sup_{\nu_\vartheta\le u<\tau\le\tau_U}\mathrm{pr}\left(X\le\boldsymbol{Z}^T\boldsymbol{\beta}_0(u)+\frac{\vartheta}{2c_2},\ X>\boldsymbol{Z}^T\boldsymbol{\beta}_0(\tau)-\frac{\vartheta}{2c_2}\right)\\&\le\sup_{\nu_\vartheta\le u\le\tau_U}\mathrm{pr}\left(\boldsymbol{Z}^T\boldsymbol{\beta}_0(u)-\frac{\vartheta}{2c_2}<X\le\boldsymbol{Z}^T\boldsymbol{\beta}_0(u)+\frac{\vartheta}{2c_2}\right)\le\vartheta/2.\end{aligned}$$

Therefore, for $n\ge\max(N_{\vartheta,\xi,1},N_{\vartheta,\xi,1})$, $\mathrm{pr}(\sup_{0<u<\tau\le\tau_U}\varphi(\widehat{\boldsymbol{\beta}}_{OSC},u,\tau)>\vartheta)\le\mathrm{pr}\big(\sup_{\tau\in[0,\tau_U]}|\phi_1\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}-\phi_1\{\boldsymbol{\beta}_0(\tau)\}|>\vartheta/8\big)+\mathrm{pr}\big(\sup_{\tau\in[\nu_\vartheta,\tau_U]}\|\widehat{\boldsymbol{\beta}}_{OSC}(\tau)-\boldsymbol{\beta}_0(\tau)\|>\frac{\vartheta}{2c_2}\big)<\xi$. This completes the proof of Lemma 3. □

Proof of Theorem 1. First note that $Y_i(t)=1-N_i(t-)-R_i(t-)$. By the definition of $\widehat{\boldsymbol{\beta}}_{PH}(\cdot)$, we have

$$\sum_{i=1}^n\boldsymbol{Z}_iN_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(\tau)\}=-\int_0^\tau\left[\sum_{i=1}^n\boldsymbol{Z}_iN_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(u-)\}\right]dH(u)+\sum_{i=1}^n\int_0^\tau\boldsymbol{Z}_i\left[1-R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(u-)\}\right]dH(u)+\boldsymbol{\nu}_n(\tau), \tag{16}$$

where $\boldsymbol{\nu}_n(\tau)$ is cadlag and satisfies that $n^{-1/2}\|\boldsymbol{\nu}_n(\tau)\|\stackrel{a}{=}0$. Viewing (16) as a stochastic integral equation for $\sum_{i=1}^n\boldsymbol{Z}_iN_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(\tau)\}$, we get from Theorem II.6.3 of [2] that

$$\sum_{i=1}^n\boldsymbol{Z}_iN_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(\tau)\}=\int_0^\tau\left(\sum_{i=1}^n\boldsymbol{Z}_i[dH(s)-R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(s)\}dH(s)]+d\boldsymbol{\nu}_n(s)\right)\boldsymbol{\pi}_{(s,\tau]}\{1-dH(u)\}.$$

Since $\boldsymbol{\pi}_{(s,\tau]}\{1-dH(u)\} = \exp\{-H(u)\,|_{u=s}^{u=\tau}\} = (1-\tau)/(1-s)$, it follows that

$$\sum_{i=1}^{n} \boldsymbol{Z}_i N_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(\tau)\} = \sum_{i=1}^{n}\boldsymbol{Z}_i\int_0^{\tau}\left[\frac{1-\tau}{1-s}\cdot\frac{ds}{1-s} - \frac{1-\tau}{(1-s)^2}R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(s)\}ds\right] + \tilde{\boldsymbol{\nu}}_n(\tau)$$
$$= \sum_{i=1}^{n}\boldsymbol{Z}_i\left[\tau-(1-\tau)\int_0^{\tau}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{PH}(s)\}}{(1-s)^2}ds\right] + \tilde{\boldsymbol{\nu}}_n(\tau), \tag{17}$$

where $n^{-1/2}\|\tilde{\boldsymbol{\nu}}_n(\tau)\| \stackrel{a}{=} 0$.

Eq. (17) implies that $n^{1/2}\boldsymbol{S}_n^{(OSC)}(\widehat{\boldsymbol{\beta}}^{(PH)}, \tau) \stackrel{a}{=} 0$. By the definition of $\widehat{\boldsymbol{\beta}}_{OSC}(\cdot)$, we also have $n^{1/2}\boldsymbol{S}_n^{(OSC)}(\widehat{\boldsymbol{\beta}}^{(OSC)}, \tau) \stackrel{a}{=} 0$ given that $\|\mathcal{G}\|$ is of order $o(n^{-1/2})$. It then follows from Lemma 2 that

$$\sup_{\tau\in[\nu,\tau_U]}\|n^{1/2}\{\widehat{\boldsymbol{\beta}}_{PH}(\tau)-\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}\| \to_p 0 \tag{18}$$

for any $\nu \in (0, \tau_U)$.

Note that $n^{1/2}\boldsymbol{S}_n^{(OSC)}(\boldsymbol{\beta}, \tau)$ has the same form that the self-consistent estimating function in (5), $n^{1/2}\boldsymbol{S}_n^{(SC)}(\boldsymbol{\beta}, \tau)$, takes except for the absence of $R_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}$ before the term $\int_0^{\tau}\frac{R_i\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(s)\}}{(1-s)^2}ds$. Therefore,

$$n^{1/2}\{\boldsymbol{S}_n^{(OSC)}(\widehat{\boldsymbol{\beta}}_{OSC}, \tau) - \boldsymbol{S}_n^{(SC)}(\widehat{\boldsymbol{\beta}}_{OSC}, \tau)\} = n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i I\{X_i > \boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(\tau)\}\int_0^{\tau}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{OSC}(s)\}}{(1-s)^2}ds.$$

Applying Lemma 3, we get $n^{1/2}\boldsymbol{S}_n^{(SC)}(\widehat{\boldsymbol{\beta}}_{OSC}, \tau) \stackrel{a}{=} 0$. By Lemma 2, this implies

$$\sup_{\tau\in[\nu,\tau_U]}\|n^{1/2}\{\widehat{\boldsymbol{\beta}}_{OSC}(\tau)-\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}\| \to_p 0 \tag{19}$$

for any $\nu \in (0, \tau_U)$. Theorem 1 then follows from (18) to (19). □

Proof of Lemma 4. Under model (1), it holds that $F_{\log T}\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)|\boldsymbol{Z}_i\} = \tau$ for $\tau \in (0, 1)$. This implies that $d\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)\}/d\tau = 1/f_{\log T}\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)|\boldsymbol{Z}_i\}$. By condition D1(ii), $f_{\log T}\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)|\boldsymbol{Z}_i\} <$ some c_3 for all i and $\tau \in (0, 1)$. Therefore, for any given $0 < \Delta < \tau_U/(2c_3)$, $|\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau) - \boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau')| \geq \Delta$ for all i and $|\tau-\tau'| \geq \Delta c_3$.

We also note that

$$\begin{aligned}
&n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)\int_0^{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\\
&\quad= n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)I\{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau) < \Delta c_3\}\int_0^{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\\
&\qquad+ n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)I\{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau) \geq \Delta c_3, \psi_i(\boldsymbol{\beta}_0,\tau) > 2\Delta c_3\}\int_{\psi_i(\boldsymbol{\beta}_0,\tau)}^{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\\
&\qquad+ n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)I\{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau) \geq \Delta c_3, \psi_i(\boldsymbol{\beta}_0,\tau) \leq 2\Delta c_3, \tau \leq 2\Delta c_3\}\int_{\psi_i(\boldsymbol{\beta}_0,\tau)}^{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\\
&\qquad+ n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)I\{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau) \geq \Delta c_3, \psi_i(\boldsymbol{\beta}_0,\tau) \leq 2\Delta c_3, \tau > 2\Delta c_3\}\int_{\psi_i(\boldsymbol{\beta}_0,\tau)}^{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\\
&\qquad+ n^{-1/2}\sum_{i=1}^{n}\boldsymbol{Z}_i R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}(1-\tau)I\{\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau) \geq \Delta c_3)\int_0^{\psi_i(\boldsymbol{\beta}_0,\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\\
&(A.2) \equiv A_{1,n}(\tau)+A_{2,n}(\tau)+A_{3,n}(\tau)+A_{4,n}(\tau)+A_{5,n}(\tau).
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
\|A_{5,n}(\tau)\| &\leq n^{-1/2}\sum_{i=1}^{n}\|\boldsymbol{Z}_i\|\int_0^{\psi_i(\boldsymbol{\beta}_0,\tau)}\frac{R_i\{\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}ds\\
&\leq c_1\int_0^{\tau}\left[n^{-1/2}\sum_{i=1}^{n}\frac{I\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(s) \leq X_i \leq \boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(s)\}}{(1-s)^2}\right]ds.
\end{aligned}$$

Using similar arguments to those for (15), we can show that

$$\sup_{\tau\in(0,\tau_U]} n^{-1/2}\sum_{i=1}^{n}\frac{I\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)\le X_i\le \boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\}}{(1-\tau)^2}\to_p 0,$$

and it follows that $\sup_{\tau\in(0,\tau_U]}\|A_{5,n}(\tau)\| = o_p(1)$.

Now it remains to show $\sup_{\tau\in(0,\tau_U]}\|\sum_{k=1}^4 A_{k,n}(\tau)\| = o_p(1)$. We first consider the situation where $\sup_{i,\tau\in[\Delta c_3,\tau_U]}|\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)-\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)|\le \Delta/2$. In this case, we can show that $\sup_{i,\tau\in(0,\tau_U]}|I\{\psi_i(\boldsymbol{\beta}_0,\tau)>2\Delta c_3,\psi_i(\widehat{\boldsymbol{\beta}},\tau)\ge\Delta c_3\}\{\psi_i(\boldsymbol{\beta}_0,\tau)-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)\}|\le\Delta c_3$. Note that $I\{\psi_i(\boldsymbol{\beta}_0,\tau)>2\Delta c_3\}=1$ implies $\tau>2\Delta c_3$. When $\boldsymbol{Z}_i^T\boldsymbol{\beta}_0\{\psi_i(\boldsymbol{\beta}_0,\tau)\}=X_i<\tau$, we have for $\tau_U>\tau\ge\psi_i(\boldsymbol{\beta}_0,\tau)+\Delta c_3$,

$$\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)\ge \boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)-\Delta/2\ge \boldsymbol{Z}_i^T\boldsymbol{\beta}_0\{\psi_i(\boldsymbol{\beta}_0,\tau)\}+\Delta-\Delta/2=X_i+\Delta/2,$$

and

$$\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}\{\psi_i(\boldsymbol{\beta}_0,\tau)-\Delta c_3\}\le \boldsymbol{Z}_i^T\boldsymbol{\beta}_0\{\psi_i(\boldsymbol{\beta}_0,\tau)-\Delta c_3\}+\Delta/2\le \boldsymbol{Z}_i^T\boldsymbol{\beta}_0\{\psi_i(\boldsymbol{\beta}_0,\tau)\}-\Delta+\Delta/2=X_i-\Delta/2.$$

This implies $|\{\psi_i(\boldsymbol{\beta}_0,\tau)-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)\}|\le\Delta c_3$. When $\psi_i(\boldsymbol{\beta}_0,\tau)=\tau<\tau_U$, we have $\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)<X_i$. Then for $u\in[\Delta c_3,\tau-\Delta c_3]$,

$$\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(u)\le \boldsymbol{Z}_i^T\boldsymbol{\beta}_0(u)+\Delta/2\le \boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)-\Delta+\Delta/2<X_i-\Delta/2.$$

This implies that either $\tau-\Delta c_3\le\psi(\widehat{\boldsymbol{\beta}}_{SC},\tau)\le\tau$ or $\psi(\widehat{\boldsymbol{\beta}}_{SC},\tau)<\Delta c_3$, and hence $|I\{\psi(\widehat{\boldsymbol{\beta}},\tau)\ge\Delta c_3\}\{\psi_i(\boldsymbol{\beta}_0,\tau)-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)\}|\le\Delta c_3$. Therefore, our claim holds, that is, $\sup_{i,\tau\in(0,\tau_U]}|I\{\psi_i(\boldsymbol{\beta}_0,\tau)>2\Delta c_3,\psi_i(\widehat{\boldsymbol{\beta}},\tau)\ge\Delta c_3\}\{\psi_i(\boldsymbol{\beta}_0,\tau)-\psi_i(\widehat{\boldsymbol{\beta}}_{SC},\tau)\}|\le\Delta c_3$.

By this result and (A.2), we get, when $\sup_{i,\ \tau\in[\Delta c_3,\tau_U]}|\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)-\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)|\le\Delta/2$,

$$\|A_{1,n}(\tau)+A_{2,n}(\tau)+A_{3,n}(\tau)+A_{4,n}(\tau)\|\le n^{-1/2}\sum_{i=1}^{n}\|\boldsymbol{Z}_i\|\cdot[\Delta c_3+\Delta c_3+2\Delta c_3+I\{X_i\le \boldsymbol{Z}_i^T\boldsymbol{\beta}_0(2\Delta c_3)\}]\equiv B_{1,n}.$$

Let $\theta(\Delta)=E(\|\boldsymbol{Z}_i\|\cdot[4\Delta c_3+I\{X_i\le\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(2\Delta c_3)\}])$ and $\sigma^2(\Delta)=\mathrm{var}(\|\boldsymbol{Z}_i\|\cdot[2\Delta c_3+I\{X_i\le\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(2\Delta c_3)\}])$. It follows from $\lim_{\Delta\to0}\mathrm{pr}(X\le\boldsymbol{Z}^T\boldsymbol{\beta}_0(\Delta c_3))=0$ that $\lim_{\Delta\to0}\theta(\Delta)=0$ and $\lim_{\Delta\to0}\sigma(\Delta)=0$. Then any $\rho>0$ and $\xi>0$, we can find $\Delta\in(0,\frac{\tau_U}{2c_3})$ such that $z_{1-\xi/3}\sigma(\Delta)+\theta(\Delta)<\rho$, where $z_{1-\xi}$ denotes the $100(1-\xi)$th percentile of a standard normal distribution. By the Central Limit Theory, for $n>$ some $N_{\rho,\xi,1}$, we have $\mathrm{pr}(B_{1,n}>z_{1-\xi/3}\sigma(\Delta)+\theta(\Delta))<2\xi/3$. Because $\sup_{\tau\in[\Delta c_3,\tau_U]}\|\widehat{\boldsymbol{\beta}}_{SC}(\tau)-\boldsymbol{\beta}_0(\tau)\|\to_p 0$ and $\|\boldsymbol{Z}_i\|<c_1$ for all i, we get $\mathrm{pr}(\sup_{i,\ \tau\in[\Delta c_3,\tau_U]}|\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)-\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)|>\Delta/2)<\xi/3$ for $n>$ some $N_{\rho,\xi,2}$. Therefore, for $n>\max(N_{\rho,\xi,1},N_{\rho,\xi,2})$,

$$\begin{aligned}\mathrm{pr}\left(\sup_{\tau\in(0,\tau_U]}\left\|\sum_{k=1}^4A_{k,n}(\tau)\right\|>\rho\right)&\le\mathrm{pr}\left(\sup_{\tau\in(0,\tau_U]}\left\|\sum_{k=1}^4A_{k,n}(\tau)\right\|>z_{1-\xi/3}\sigma(\Delta)+\theta(\Delta)\right)\\&\le\mathrm{pr}\left(\sup_{i,\tau\in(0,\tau_U]}|\boldsymbol{Z}_i^T\widehat{\boldsymbol{\beta}}_{SC}(\tau)-\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)|>\Delta/2\right)\\&\quad+\mathrm{pr}(B_{1,n}>z_{1-\xi/3}\sigma(\Delta)+\theta(\Delta))<\xi.\end{aligned}$$

This proves $\sup_{\tau\in(0,\tau_U]}\|\sum_{k=1}^4A_{k,n}(\tau)\|\to_p 0$ and thus completes the proof for Lemma 4. □

References

[1] O.O. Aalen, Nonparametric inference for a family of counting processes, Annals of Statistics 6 (1978) 701–726.
[2] P.K. Andersen, E. Borgan, R.D. Gill, N. Keiding, Statistical Models Based on Counting Processes, second ed., Springer-Verlag, New York, 1998.
[3] M. Buchinsky, J. Hahn, A alternative estimator for censored quantile regression, Econometrica 66 (1998) 653–671.
[4] B. Efron, 1967, The two-sample problem with censored data, Proc Fifth Berkley Symposium in Mathematical Statistics, IV, pp. 831–553.
[5] B. Fitzenberger, A guide to censored quantile regressions, Handbooks of Statistics: Robust Inference 15 (1997) 405–437.
[6] M. Fygenson, Y. Ritov, Monotone estimating equations for censored data, The Annals of Statistics 22 (1994) 732–746.
[7] B. Honore, S. Khan, J. Powell, Quantile regression under random censoring, Journal of Econometrics 109 (2002) 67–105.
[8] Y. Huang, Quantile calculus and censored regression, The Annals of Statistics 38 (3) (2010) 1607–1637.
[9] E. Kaplan, P. Meier, Nonparametric estimation from incomplete observations, Journal of the American Statistical Association 53 (1958) 457–481.
[10] R. Koenker, Censored quantile regression redux, Journal of Statistical Software 27 (2008) http://www.jstatsoft.com.
[11] R. Koenker, 2008, Quantreg: Quantile Regression, R package version 4.16, http://www.r-project.org.
[12] R. Koenker, G. Bassett, Regression quantiles, Econometrica 46 (1978) 33–50.
[13] I. McKeague, S. Subramanian, Y. Sun, Median regression and the missing information principle, Journal of Nonparametric Statistics 13 (2001) 709–727.
[14] W. Nelson, Theory and applications of hazard plotting for censored survival data, Technometrics 14 (1972) 945–965.
[15] T. Neocleous, S. Portnoy, On monotonicity of regression quantile functions, Statistics and Probability Letters 78 (2008) 1226–1229.

[16] T. Neocleous, K. Vanden Branden, S. Portnoy, Correction to censored regression quantiles by S. Portnoy, Journal of the American Statistical Association 98 (2003) 1001–1012. 101, 860–861.
[17] L. Peng, Y. Huang, Survival analysis with quantile regression models, Journal of the American Statistical Association 103 (2008) 637–649.
[18] S. Portnoy, Censored regression quantiles, Journal of the American Statistical Association 98 (2003) 1001–1012.
[19] S. Portnoy, G. Lin, Asymptotics for censored regression quantiles, Journal of Nonparametric Statistics 22 (2010) 115–130.
[20] J. Powell, Least absoluate deviations estimation for the censored regression model, Journal of Econometrics 25 (1984) 303–325.
[21] J. Powell, Censored regression quantiles, Journal of Econometrics 32 (1986) 143–155.
[22] B.W. Turnbull, Nonparametric estimation of a survivorship function with double censored data, Journal of the American Statistical Association 69 (1974) 169–173.
[23] B.W. Turnbull, The empirical distribution function with arbitrarily grouped, censored and truncated data, Journal of the Royal Statistical Society B 38 (1976) 290–295.
[24] H. Wang, L. Wang, Locally weighted censored quantile regression, Journal of the American Statistical Association 104 (2009) 1117–1128.
[25] Z. Ying, S.H. Jung, L.J. Wei, Survival analysis with median regression models, Journal of the American Statistical Association 90 (1995) 178–184.